PRE-EXISTING DETERMINANTS OF APPRAISAL INTERVIEW SUCCESS

by

MATT L. RIGGS

B.A., Bethany College, 1979

A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1986

Approved by:

Major Professor

TABLE OF CONTENTS

LIST OF TABLES

Previous investigations into the performance appraisal process have generally been based upon the assumption that the performance appraisal interview (in which supervisor and subordinate discuss past performance and plan improvements in future work) must be perceived as appropriate by the subordinate to have the desired effects (i.e., increased motivation, improved performance, enhanced supervisor-subordinate relations, etc.). The main objective of most studies to this point has been the identification of specific characteristics of the appraisal process itself that might lead to the acceptance and consequent utilization of the evaluation by the employee. Events occurring within the interview found to be correlated with its success have been classified into six general categories (Burke, Weitzel and Weir, 1978): 1) the level of subordinate participation in the appraisal process, 2) the "helpfulness" and/or "constructiveness", (and/or "supportiveness") of the supervisor's attitude, 3) the degree to which specific goals to be obtained by the subordinate are mutually established by the supervisor and subordinate, 4) the extent to which problems are addressed and perceived to be solved, 5) the amount of criticism given by the supervisor, and 6) the proportion of time that the subordinate speaks.

More recent research has investigated other potential determinants of appraisal success. The amount of threat perceived by the subordinate during the evaluation was found to be negatively related to satisfaction with the interview and desire to improve future performance (Burke, Weitzel and Weir, 1980). Smircich and Chesser (1981) investigated the "level of

authenticity" in the interaction that occurs during evaluation sessions. They found no support for their proposition that subordinate acceptance of appraisals would be correlated with the subordinate's perception that the interview went beyond formal roles, game playing and superficial courtesies.

One limitation to the research cited above is the attempt to explain the appraisal interview outcomes solely by the events that occur within the context of the interview itself. It is reasonable to assume that subordinates bring into the appraisal interview situation a good deal of cognitive and emotional "baggage" that may dramatically influence their perceptions of and reactions to whatever events occur during the interview itself. Although actual research designed to identify critical pre-evaluative conditions is rare, awareness of their importance has been demonstrated by some authors.

Lawler (1967) suggested that the success or failure of the evaluation process is dependent upon more than just the objective characteristics of the rating system itself. Although Lawler conducted no research on this issue, he proposed that individual personality differences among raters and ratees (e.g., need for feedback and authoritarianism), differences in organizational characteristics (e.g., Theory X and Theory Y organizations) and attitudes of the raters toward the fairness and acceptability of the rating system might all contribute to the perceived validity and consequent effectiveness of the evaluative process.

Shrauger (1975), Shrauger and Lund (1975) and Baird (1977) suggested that self-perceptions and self-esteem levels of the

individuals being evaluated may significantly affect their reaction to the process. These studies suggested that subordinates with high self-perceptions or high self-esteem would react differently to the evaluation and the evaluators depending on the nature of the feedback (they would agree with and accept positive feedback, but disagree with and reject negative), while the perceptions of those with low self esteem did not vary in response to the nature of the evaluation (both good and bad feedback were accepted more analytically and objectively with little or no apparent effect on perception of or agreement with the supervisor). Differential ego-involvement is given as a possible reason for these contrasts. More specifically, high self-esteem subjects may have been more ego-involved to protect their opinions of themselves while low self-esteem subjects may have had nothing to lose.

In his discussion concerning the above research, Shrauger (1975) speculated about the cognitive processing that might be occurring in the subordinate's response to the appraisal. He proposed that in any evaluative process, the individual being evaluated will assess the credibility of the judgement source. In a work setting, this implies that subordinates will attempt to decide whether the supervisor has the appropriate information about them, as well as the expertise to make a valid assessment. This judgement will most certainly hinge upon experiences of the subordinates that have occurred prior to the initiation of the appraisal interview.

Shrauger's (1975) postulates were consistent with research findings by Knight and Weiss (1980). They found that not only

were new leaders perceived to be more competent by a lab group when the person responsible for the selection of the leader was perceived to be competent, but also that the leaders were significantly more influential when perceived by the group to be more competent. In a similar study, Knight and Saal (1984) replicated the results of perceived competence of the leader as an effect of the perceived competence of the selection agent. They also demonstrated that the credibility of the leader was assessed by group members according to the gender-type of the groups' task (leaders of groups with male-oriented tasks were perceived to be more competent than those of female-oriented tasks). Also, in feminine task conditions, female leaders or those chosen by female agents were perceived as more competent than male leaders or those chosen by male agents. All of these effects demonstrated the tendency of individuals to evaluate the ability of immediate supervisors on the basis of situational factors external to the actual supervisor-subordinate interaction.

Landy, Barnes and Murphy (1978) mailed questionnaires to 950 production line employees concerning the nature of, and their satisfaction with, their appraisal interviews. They found that the frequency of previous evaluation and the subordinate's perception of the supervisor's knowledge were both predictive of subsequent subordinate satisfaction that performance had been fairly and accurately evaluated. In their discussion, Landy et al. suggested that these results might indicate that supervisors could improve the effectiveness of their work

performance evaluations by devoting sufficient time to observation and evaluation of the subordinate's performance prior to the evaluative session. In addition to concerns for the subordinate's perception of the validity of the evaluation information, spending sufficient time in direct observation may have practical implications for the validity of the work assessments. Henneman and Wexley (1983) demonstrated that performance ratings did increase in accuracy as observation of the ratee's performance increased.

Ilgen, Peterson, Martin and Boeschen (1981) also concluded that evaluation sessions do "not occur in a vacuum", but reflect supervisor-subordinate interactions that preceed the appraisal session. They surveyed one hundred and six supervisor-subordinate pairs in a wood-products industry to obtain measures of both independent and dependent variables. Their field study indicated that a significant correlation existed between the subordinate's perceptions that feedback had occurred with a high frequency prior to the appraisal interview and the subordinate's consequent satisfaction with the appraisal session in terms of its positive atmosphere, helpfulness, and specificity of information.

SHORTCOMINGS OF PREVIOUS RESEARCH

Response Bias

The predominant method of research in previous studies involved collecting subordinates' subjective perceptions of the appraisal process itself (via survey questionnaires) and correlating these perceived characteristics with measures of the subordinates' satisfaction with and reaction to the appraisal

(also obtained by survey items). Typical measures of interview characteristics included questions such as, "Think of your last performance review session. How well does this statement describe it?", 'My supervisor was helpful and constructive.' (Burke et al., 1978) or, "I am given the opportunity to state 'my side' of the issue" (Greller, 1975). Responses were generally scored on Likert scales. Dependent measures obtained in the same manner included such items as, "In my opinion, the session went very well" (Ilgen et al., 1981), and "How fair do you feel your last performance appraisal session was?" (Burke et al., 1978).

A great potential for the occurrance of response bias is created by the use of such survey data from the same subject pool as both independent and dependent variables. Rather than being measures of true relationships between independently determined variables, correlations obtained by this method may simply be the product of the subjects' preconceptions concerning the expected results of certain evaluative environments. "Knowing" that higher levels of subordinate involvement result in "better" evaluations would cause subjects to respond to items in a manner consistent with this "knowledge", resulting in a correlation between these events in the survey data--whether or not such a relationship actually existed.

Accuracy of Subject Observations

A serious question must be raised concerning the effectiveness of using subordinate perceptions as exclusive indicators of what occurs in appraisal interviews. Supervisors' perceptions of the characteristics and consequent value of the

appraisal interviews have generally been found to differ significantly from those of the subordinates (Burke, Weitzel & Deszca, 1982; Burke et al., 1980; Ilgen et al., 1981; Smircich & Chesser, 1981). Supervisors have generally been found to be more positive in both their perceptions of the nature of the appraisal interview itself (e.g., much praise was used, the atmosphere was helpful and positive, the subordinate participated a great deal, etc.) and in their observations of the interview's outcomes (e.g., the interview resulted in increased performance, job-related problems were identified and solved, etc.).

The important question raised by inconsistency in subordinate and supervisor perceptions is not whose perception of reality is correct. The issue is whether participants in the interview process can describe its true nature and outcomes accurately enough to enable the determination of appraisal characteristics that may be related to various criteria. If the measures of the independent and dependent variables were based upon innacurate observations, the researcher's subsequent conclusions concerning what conditions make for a good appraisal interview may be unfounded.

It is therefore suggested that a lab approach should be applied to determine what pre-existing factors affect success in appraisal interviews. This would eliminate the problem of accurately determining the true nature of the treatment received. Rather than depending upon the subject's perceptions and descriptions of the independent variables, a lab design with controlled settings would enable the investigator to determine the true nature of the appraisal. Objective observation of

relevant outcomes by the investigator within the lab setting also removes the opportunity for the subjects to bias the measurement of the dependent variables.

Limited Dependent Measures

In addition to the variables of the subordinates' subjective perceptions of and consequent attitudes due to the appraisal experiences, a comprehensive study should include more objective dependent measures. Behavioral observations of desired responses and simple task performance measures are dependent measures that should be of more immediate interest to production-oriented organizations. More specifically, typical outcomes used in previous research such as the subject's perception of "greater mutual understanding", "satisfaction with the appraisal", or "expected effects on job performance" do give some insight to underlying processes. Nevertheless, dependent measures such as observation of the subordinate's implementation of appraisal directives or direct performance measures indicating faster and/or better production should be of more concern in applied settings. By themselves, good intentions are of little value, and it may be wrong to assume that the previously measured attitudinal states were necessarily indicative of desirable levels of behavior and performance.

Reliance upon Correlational Analysis

Previous research has relied exclusively upon correlational analysis. Caution must be used in the interpretation of any correlational study. The existence of a relationship between certain perceived appraisal conditions and subjectively described

appraisal outcomes does not tell anything about causal relationships. Landy et al. (1978) warned that pre-existing positive attitudes toward appraisals and/or supervisors may result in preferential treatment from supervisors and different evaluative experiences--real or imagined. In other words, the positive attitude may not be the result of the type of appraisal experienced, but the direct cause of the type of appraisal experienced.

RESEARCH PURPOSE

Events that occur within the appraisal setting undoubtedly have an effect upon the success of the interview. They have received the bulk of research attention and--despite noted shortcomings in the investigative procedures--are considered by some to be well documented. It has been proposed, however, that the success of an appraisal interview may be largely determined before the subordinate ever sets foot in the supervisor's office. If this is true, it is important for the supervisor to know what steps might be taken prior to the appraisal to increase the chances for a successful outcome. Although contributing factors to a successful interview such as personality variables and organizational characteristics (Lawler, 1967) or subordinate self-esteem levels (Baird, 1977; Shrauger, 1975; Shrauger & Lund, 1975) may not be within the direct control of the supervisor, his/her own credibility to the subordinate (as posited by Shrauger, 1975) might be enhanced through personal endeavor. The findings of Ilgen et al. (1981) concerning the nature of previous interactions, and the results of Landy et al. (1978) with frequency of evaluation and perceived supervisor

knowledge lend support to the proposition that effectiveness of appraisal interviews might be increased by maintaining consistent contact with the subordinate prior to the appraisal interview itself and by successfully demonstrating personal competence at the task being evaluated.

It was therefore suggested that, although other contributing factors may exist, two crucial variables that may affect the subordinate's appraisal attitudes are (1) the degree to which the subordinate perceives that the supervisor is aware of his/her actual work performance, and (2) the amount of "expert knowledge" that the subordinate believes the supervisor is bringing into the interaction. It was assumed that credibility of any information conveyed to the subordinate by the supervisor is logically dependent upon these two variables, and implementation of evaluative information should be dependent upon the information's perceived credibility.

It was further proposed that a lab test of the above propositions would be superior to previous investigations in its ability to determine whether levels of supervisory observation and perceived supervisor knowledge have causal relationships with levels of subordinate acceptance and implementation of evaluative information, and subsequent levels of subordinate performance. Controlled determination of the independent variables, objective observation of multiple dependent variables, and analysis that enables testing of causation rather than interpretation of associations are all advantages to be gained by such an approach.

HYPOTHESES

1. Observation of the subjects' performance prior to the appraisal will enhance the subjects' subjective reactions to and compliance with the evaluative appraisal, as well as their subsequent performance of the same task. In the close observation condition, 1) the subordinate will perceive the appraisal to be fairer and more accurate, 2) the subordinate's evaluation of the value of the appraisal suggestions will be more positive, 3) the subordinate's compliance with the evaluative suggestions will be greater, and 4) the subordinate's performance of work following the evaluation will be faster and more accurate.

2. Increased supervisory competence will enhance the subjects' subjective reactions to and compliance with the evaluative appraisal, as well as their subsequent performance of the same task. In the "competent" condition, 1) the subordinate will perceive the appraisal to be fairer and more accurate, 2) the subordinate's evaluation of the value of the appraisal will be more positive, 3) the subordinate's compliance with the evaluative suggestions will be greater, and 4) the subordinate's performance of work following the evaluation will be faster and more accurate.

METHOD

SUBJECTS

Subjects were volunteers from undergraduate psychology classes. Eighty subjects were randomly assigned to one of the four experimental conditions, resulting in twenty subjects per cell. Thirty-six subjects were male, and forty-four were female.

DESIGN

A 2x2 factorial design with three categories of dependent measures was used. Factors included: (1) level of supervisory observation of task (direct observation or no observation) and (2) level of expressed supervisory competence (very competent or not competent).

Dependent measures were: 1) responses to a post-experiment questionnaire designed to determine perceptions of the fairness and accuracy of the appraisal session and the value of the appraisal recommendations, 2) an objective scoring of the subject's compliance to suggestions made during the appraisal session, and 3) two performance measures that consisted of time and accuracy scores from the first and second trials.

PROCEDURE

The subject was seated at a microcomputer console. The experimenter then started a videotape containing an introduction of himself, orientation information, and directions for the initial task. The experimenter was described as either: 1) an experienced microcomputer word processing specialist, listing impressive experience and credentials (see stimulus materials for exact introductions) or 2) a psychology graduate student with little experience or knowledge in word processing who was running the experiment for someone else (see stimulus materials section). The subject was told that he or she was involved in an experiment designed to test his or her ability to learn and use a new word processing program.

The video then instructed the subject how to perform the

word processing task that he or she would be asked to complete. The subject was taught four functions: cursor movement, text insertion, text deletion, and text revision (changing one letter to another). Along with these audio-visual instructions, the subject received a reference sheet that fit around the keyboard of the microcomputer containing identical instructions. Additional functions (to be introduced later, following the first trial) were given on another reference sheet concealed under the first sheet. The experimenter then loaded a short passage of text containing fifteen errors into the microcomputer and gave the subject a sheet of instructions explaining what corrections were to be made. In the "expert" condition, the experimenter loaded the text without error. In the "non-expert" condition, the experimenter initially inserted the floppy disk incorrectly and had to repeat the procedure before the first trial could begin.

As the subject performed the task for the first time, the experimenter did one of the following: (1) directly observed the subject as he/she performed the task, or (2) left the room, returning when the subject was done with the task (the experimenter instructed the subject to signal when the task was completed). Performance of the task was recorded on a video recorder, using in-line recording of the subject's computer entries (i.e, the computer's video output was directed both to the subject's television monitor and to the video recorder). This recording was made without the subject's knowledge.

Upon completion, the experimenter delivered an appraisal of the subject's performance (identical for all subjects) during

which comments were made about the performance on the first trial of the task and directions for the second trial were given. Information about ways to perform more efficiently were made available to the subject during this session. Some of this information was actually designed and intended to improve task performance--some of it was apparently irrelevant to the task. The suggestions lacking in face validity were intended to test the subject's willingness to follow evaluation directives. Immediately following the appraisal, a copy of the subject's work was printed out for later scoring.

Included in the experimenter's suggestions were alternate functions that the subject could choose to practice and then use during the second trial. The subject was directed to the instructions for these additional functions located underneath the first reference sheet. The experimenter then announced that the second trial would begin in 10 minutes. The screen was left as it was at the start of the subject's first trial (containing a short text with fifteen errors). The subject was left with the option to practice or rest as he/she desired.

The subject's subsequent use of the time was recorded by the video recorder (which was still in operation). The experimenter then returned, started the subject on the second task, and then sat at a desk at the back of the room as the subject performed a task parallel to the one performed in the first trial. This trial was also videotaped and later scored for the number of behaviors that were in compliance with the directives of the evaluation (see dependent measures section). Upon completion of the task,

the experimenter gave the subject a short survey of his/her impression of the evaluative session. This survey was completed while the experimenter printed out the second version of the work task. Print outs were later scored for their accuracy, and videotapes of the trials were timed to determine completion times.

STIMULUS MATERIALS

The following introductions were included in the video shown to the subject after he/she entered the room:

> Thank you for volunteering to participate in this study. My name is Matt Riggs.
>
> VERSION A (competent): I am currently a graduate student in KSU's industrial psychology program, but having previously worked for three years as a district representative for CBS software systems specializing in word processing, I have received some interest from my former employers in doing research into the ability of people to learn word processing programs.
>
> VERSION B (not competent): I am currently a graduate student in KSU's industrial psychology program, and I am conducting this research for my major advisor, Dr. Knight. I mention this due to the fact that any questions you may have concerning the word processing system we are using would be best directed to him as my experience with the process is limited to what I need to know to load the program and print out your finished work.
>
> The purpose of this research is to measure the

average individual's ability to learn and use the skills needed to perform word processing tasks with a limited amount of time and training. You will receive an instruction sheet for some of the program's functions and be shown a video training tape of how they are to be used. Keeping the instruction sheet, you will then attempt to use the program to edit and correct a short passage of text that I will load onto the machine for you.

Following the first trial, I will evaluate your performance, making suggestions for improvement. You will then perform a similar task in a second trial. Following this last trial, I will ask you to fill out a short survey of your reactions to the experience.

I foresee no risk or discomfort to you in this experiment. The benefits to you will not only include some limited exposure to this particular word processing program, but also some insight into methods and procedures used by psychologists to measure human performance.

Before the initial video demonstration of the four processing functions, the subjects received a reference sheet to be used as they wished during the task. During the audio-video demonstration, a voice track of the experimenter reading these instructions accompanied a visual display of what the subjects would see on the screen when they correctly completed each procedure. The instructions appeared as follows:

CURSOR MOVEMENT

The CRSR key moves the cursor to the right when the "shift" key is up , to the left when the "shift" key is held down.

TEXT INSERTION

Move the cursor to the space where the word is to be inserted, hold down the shift key and press the INST/DEL key to create as many spaces as necessary. Type the word to be inserted into the newly created space.

CHANGING TEXT

Place the cursor under the incorrect letter and type the correct letter.

TEXT DELETION

Move the cursor to the start of the text to be deleted, hold down the CTRL key, and press the <-- key until the desired amount of text has been removed.

Under the first reference sheet was a second describing the functions suggested by the experimenter during the appraisal session. The following information appeared on the second sheet:

CURSOR MOVEMENT

To move the cursor ahead one whole word at a time, press the f1 key.

TEXT INSERTION

To insert a word where no space exists, place the cursor under the space where you wish the word to begin, press the CTRL key and hold it down while you press the I (for insert) key. Now type the word or words you wish to insert. IMPORTANT--After inserting the word, exit insert mode by once again holding down the CTRL key while pressing the I key.

TEXT DELETION

To erase entire words at once, place the cursor under the first letter of the word or words you wish to delete, press the CTRL key, and hold it down while you press the E (for erase) key. Now press the W key to erase as many words as you wish. IMPORTANT--After erasing, exit erase mode by pushing the RETURN key.

REVERSING TWO LETTERS

To reverse the order of two letters, place the cursor under the first of the two letters to be switched, press down the CTRL key and hold it down as you press the X key.

One of two texts (see Appendix A) were then presented to the subjects on the screen of the word processor (the order of presentation was randomly assigned to avoid order effects). Subjects were also given the appropriate sheet of directives that described the corrections to be made. Corrections were listed in groups according to the type of correction to be made (see Appendix B).

The evaluation stimulus was as follows, spoken by the experimenter:

"O.K., from what I saw of your first try, I think you can improve your performance for the second trial. I'd like you to make the corrections faster while maintaining or decreasing your number of errors. Let me make some suggestions."

"It appeared as though you had some trouble keeping track of what corrections to make next and where to make them. It might be helpful for you to complete the corrections in the order in

which they appear in the text rather than the order in which they appear on your direction sheet. Check the copy of the completed text for the correct order. I think doing it that way might help."

"Also, to make sure you don't overlook corrections, I suggest that you make a checkmark beside each item on your direction sheet as the correction is completed."

"I also think you should try some additional functions. You appeared comfortable with the functions being used in the first trial, so I think you can easily learn some new ones. They might speed things up a bit. Here are some other ways of doing what you have already been doing. A little practice with these would improve your performance."

"I'll leave the screen as it was at the start of the first trial. I'll be back in about ten minutes. We'll start the second trial then. In the meantime, I think it would be a good idea for you to practice, but you can just rest if you'd like."

DEPENDENT MEASURES

Upon completion of the second trial, the subject was asked to complete a questionnaire (see Appendix C) that included items intended: to measure his or her perceptions of the accuracy and fairness of the appraisal (item 8), to measure his or her attitudes concerning the value of the appraisal's content (items 5 and 10), and to check manipulations of the independent variables (items 1 and 7).

The following were used as measures of the subject's compliance to the appraisal session's directives. Each was used as an isolated measure of compliance. The purpose of these

measures were to determine the actual behavioral responses of the subjects to the appraisal directives within the different treatment conditions.

A. Subject's use of new functions: the experimenter counted the number of corrections in which the new functions given during the appraisal session were used. Separate scores were kept for new functions used during the practice session and new functions utilized during the second task. Corrections were scored from the videotape of the subjects' computer entries. One point was scored every time a new function was attempted (whether it was used correctly or not).

B. The following was determined by the experimenter by looking at the worksheet used by the subject after the task was completed or by observing the videotape of the subject's performance of the second trial.

___ a. the subject made a checkmark beside each task on the direction sheet. Score one point if the subject did as directed, score two points if the subject did as directed at least half, but not all of the time, score three points if the subject did as directed less than half of the time.

___ b. the subject completed the corrections in the order in which he/she was directed. One

point was scored for each function completed in the correct order.

The two "performance scores" were the following raw scores: the number of errors made on the second trial of the task (determined by scoring the printouts of the second task trials), and the amount of time used to complete the second trials (obtained by timing the videotape records). Errors were scored on the basis of whether the correct letters and words appeared in the text. Variations in spacing from the example text were also counted as errors. The purpose of these measures was to determine whether different treatments might result in different levels of ultimate performance during the second task.

RESULTS

Reliability Coefficients

Student scorers observed videotapes of the subjects' entries to quantify the dependent measures of practice, movement, order, and corrections. The reliability of these scores was established by having six tapes scored by both scorers. Correlations of .98 to .99 were obtained for these four dependent measures.

Student scorers also tabulated the dependent measure of errors by counting errors in each of the subjects' Trial 1 and Trial 2 printouts. Reliability of this scoring was once again obtained by having both scorers count errors on six printouts. The correlation of their scores for this measure was .98.

Manipulation Checks

A manipulation check for the independent variable of supervisor's expertise was based upon an ANOVA of the subjects' responses to question number 1 on the survey. Results are shown

in Tables 1 and 2. This manipulation was very effective, evidenced by a significant main effect of experimenter expertise, $F(1,76) = 77.949$, $p < .001$. No other effects approached significance.

The manipulation check for the independent condition of supervisory observation level was not so successful. An ANOVA of responses to item number 7 was intended to measure the effects of the supervisor observing the first trial or leaving the room during the first trial. Results for item number 7 are shown in Tables 3 and 4. Main effects for level of observation failed to reach significance. Instead, a significant main effect for supervisor's expertise occurred ($F(1,76) = 12.032$, $p < .01$). A significant interaction also occurred for item number 7 ($F(1,76) = 5.563$, $p < .05$). In the non-expert condition, the subjects perceived more accurate observation when they had been observed (as expected). In the expert condition, the supervisor received approximately the same scores in both observation conditions (contrary to expectations).

A one-way analysis of variance was conducted for item number 7 ($F(3,76) = 6.949$, $p < .001$). Student-Newman-Keuls post-hoc analysis of significant differences between cell means for item number 7 revealed that the means in the expert/observed, non-expert/observed, and expert/non-observed conditions were all significantly higher than the mean of the non-expert/non-observed condition.

MANOVAs

Multiple analysis of variance was completed for groups of

TABLE 1

Analysis of Variance of Subjects' Perception of the Supervisor's Expertise

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	46.512	77.949***	0.50
LEVEL OF OBSERVATION	1	0.012	0.021	--
EXPERTISE X OBSERVATION	1	0.012	0.021	--
RESIDUAL	76	0.597		
TOTAL	79	1.163		

*** p < .001

TABLE 2

Mean Values and Standard Deviations of Subjects' Perceptions of Supervisor's Expertise

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	1.550 (.605)	1.600 (.754)	1.575 (.675)
NON-EXPERT	3.100 (.718)	3.100 (.968)	3.100 (.708)
TOTAL	2.325 (1.023)	2.350 (1.145)	2.338 (1.079)

NOTE: ( ) = std.dev.

Scale Values: Low values indicate highest perception of expertise.

TABLE 3

Analysis of Variance of Subjects' Perception of the Accuracy of the Supervisor's Observation of the First Trial

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	7.812	12.032**	0.11
LEVEL OF OBSERVATION	1	2.112	3.253	0.02
EXPERTISE X OBSERVATION	1	3.612	5.563*	0.04
RESIDUAL	76	0.649		
TOTAL	79	0.796		

** p < .01
* p < .05

TABLE 4

Mean Values and Standard Deviations of Subjects' Perceptions of Accuracy of Supervisor's Observations

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	1.700 a (.657)	1.600 a (.598)	1.650 (.622)
NON-EXPERT	1.900 a (.641)	2.650 b (1.182)	2.275 (1.012)
TOTAL	1.800 (.649)	2.125 (1.067)	1.963 (.892)

( ) = std. dev.
Scale Values: Low values indicate perception of accurate observation.
a,b: Cell means with common indices are not significantly different by the Student-Newman-Keuls procedure, p < .05.

dependent measures designed to measure: 1) subject attitudes toward the appraisal, 2) the behavioral compliance of the subjects to the supervisor's instructions, and 3) the performance of the subjects in the second trial. Results of the MANOVA for the three attitudinal variables (item numbers 5, 8 and 10 on the survey) are shown in Table 5. With the use of Wilks' criterion, a significant effect for level of expertise was obtained, $F(3,74) = 3.072$, $p < .05$.

Results of the MANOVA for the five compliance measures are shown in Table 6. Again according to Wilks' criterion, significant effects were obtained for the main effect of level of observation, $F(5,72) = 2.407$, $p < .05$, and the interaction of observation and expertise, $F(5,72) = 4.224$, $p < .01$.

The MANOVA for performance measures (time and errors) failed to reach significance (see Table 7, means and standard deviations in appendix D). Univariate ANOVA's were consequently completed only for the attitudinal and compliance measures.

ANOVAs

Results of the univariate analyses of variance for the attitudinal measures are shown in Tables 8 through 13. The survey item that measured the subjects' perception of the value of the supervisor's suggestions (item number 5, Tables 8 and 9) resulted in a significant main effect for level of expertise, $F(1,76) = 4.468$, $p < .05$. Measurement of the subjects' perception of the value of the appraisal for improving performance (item #10, Tables 10 and 11) also produced a main effect for expertise, $F(1,76) = 8.812$, $p < .01$. Measurement of the subjects' perception of the value of the evaluation in a job context (item

TABLE 5

Multivariate Analysis of Variance of Attitudinal Measures

SOURCE	df	Wilks' Lambda	Approx F	Eigen.
SUPERVISOR'S EXPERTISE	3	0.889	3.072*	0.13
LEVEL OF OBSERVATION	3	0.956	1.142	0.05
EXPERTISE X OBSERVATION	3	0.916	2.256	0.09
RESIDUAL	74			
TOTAL	79			

* p < .05

TABLE 6

Multivariate Analysis of Variance of Compliance Measures

SOURCE	df	Wilks' Lambda	Approx F	Eigen.
SUPERVISOR'S EXPERTISE	5	0.898	1.620	0.11
LEVEL OF OBSERVATION	5	0.857	2.407*	0.17
EXPERTISE X OBSERVATION	5	0.773	4.224**	0.29
RESIDUAL	72			
TOTAL	79			

* p < .05
** p < .01

TABLE 7

Multivariate Analysis of Variance of Performance Measures

SOURCE	df	Wilks' Lambda	Approx F	Eigen.
SUPERVISOR'S EXPERTISE	2	0.988	0.439	0.01
LEVEL OF OBSERVATION	2	0.993	0.266	0.01
EXPERTISE X OBSERVATION	2	0.959	1.592	0.04
RESIDUAL	75			
TOTAL	79			

TABLE 8

Analysis of Variance of Subjects' Perception of the Value of the Supervisor's Suggestions

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	3.612	4.468*	0.04
LEVEL OF OBSERVATION	1	1.012	1.252	< .01
EXPERTISE X OBSERVATION	1	0.312	0.386	--
RESIDUAL	76	0.809		
TOTAL	79	0.840		

* p < .05

TABLE 9

Mean Values and Standard Deviations of Subjects' Perceptions of Value of Suggestions

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	2.050 (.999)	1.700 (.571)	1.875 (.822)
NON-EXPERT	2.350 (.933)	2.250 (1.020)	2.300 (.966)
TOTAL	2.200 (.966)	1.975 (.862)	2.088 (.917)

( ) = std. dev.
Scale Values: Low values indicate suggestions highly effective

TABLE 10

Analysis of Variance of Subjects' Perception of the Value of the Appraisal for Improving Performance

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	7.200	8.812**	0.09
LEVEL OF OBSERVATION	1	0.050	0.061	--
EXPERTISE X OBSERVATION	1	2.450	2.998	0.02
RESIDUAL	76	0.817		
TOTAL	79	0.909		

** p < .01

TABLE 11

Mean Values and Standard Deviations of Subjects' Perceptions of Appraisal's Value to Performance

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	1.850 (.988)	1.450 (.510)	1.650 (.802)
NON-EXPERT	2.100 (.852)	2.400 (1.143)	2.250 (1.006)
TOTAL	1.975 (.920)	1.925 (.997)	1.950 (.953)

Scale Values: Low values indicate suggestions improved performance.

TABLE 12

Analysis of Variance of Subjects' Perception of the Value of the Evaluation in a Job Context

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	3.612	4.150*	0.04
LEVEL OF OBSERVATION	1	0.612	0.704	--
EXPERTISE X OBSERVATION	1	4.512	5.184*	0.05
RESIDUAL	76	0.870		
TOTAL	79	0.948		

* p < .05

TABLE 13

Mean Values and Standard Deviations of Subjects' Perceptions of the Evaluation in a Job Context

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	2.100 a,b (.852)	1.800 a (1.005)	1.950 (.932)
NON-EXPERT	2.050 a,b (.945)	2.700 b (.923)	2.375 (.979)
TOTAL	2.075 (.888)	2.250 (1.056)	2.163 (.974)

Scale Values: Low values indicate appraisal was fair and accurate.
a,b: Cell means with common indices are not significantly different by the Student-Newman-Keuls procedure, p < .05.

number 8, Tables 12 and 13) resulted in a third significant main effect for expertise, $F(1,76) = 4.150$, $p < .05$. Analysis of number 8 also resulted in a significant interaction effect ($F(1,76) = 5.184$, $p < .05$).

A one-way analysis of variance was conducted for item number 8 ($F(3,76) = 3.346$, $p < .05$). Student-Newman-Keuls analysis revealed that the only significant difference occurred between the means for expert and non-expert supervision in the non-observed condition. The expert/non-observed condition produced a higher rating than did the non-expert/non-observed condition. Though this effect is described, it should be interpretted with caution. This interaction is not supported by the established criterion of a significant MANOVA effect.

Results for the univariate ANOVAs for compliance measures are shown in Tables 14 through 23. Three of the five measures resulted in significant main effects for level of observation. The results for the use of new functions during the practice period ($F(1,76) = 5.456$, $p < .05$), the use of the prescribed movement function during the second trial ($F(1,76) = 5.094$, $p < .05$), and the use of new functions to make corrections during the second trial ($F(1,76) = 6.099$, $p < .05$) all indicated that level of observation positively influenced the subjects' willingness to implement the supervisor's directives.

Measurement of the use of the prescribed movement function (see Table 16) also resulted in a strong interaction, $F(1,76) = 15.277$, $p< .001$. One-way analysis of variance was conducted for this variable ($F(3,76) = 7.563$, $p < .001$). Student-Newman-Keuls post-hoc analysis indicated that the cell mean from the non-

TABLE 14

Analysis of Variance of Subjects' Use of New Functions During the Practice Period

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	8.450	0.476	--
LEVEL OF OBSERVATION	1	96.800	5.456*	0.05
EXPERTISE X OBSERVATION	1	0.450	0.025	--
RESIDUAL	76	17.743		
TOTAL	79	18.407		

* p < .05

TABLE 15

Mean and Standard Deviation of Number of New Functions Used by Subjects During the Practice Period

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	4.350 (5.264)	2.000 (4.243)	3.175 (4.867)
NON-EXPERT	3.550 (4.419)	1.500 (2.395)	2.525 (3.658)
TOTAL	3.950 (4.814)	1.750 (3.410)	2.850 (4.290)

Scale Values: Low values indicate less use of new functions.

TABLE 16

Analysis of Variance of Subjects' Use of the Prescribed Movement Function During the Second Trial

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	86.112	2.320	0.01
LEVEL OF OBSERVATION	1	189.112	5.094*	0.04
EXPERTISE X OBSERVATION	1	567.112	15.277***	0.14
RESIDUAL	76	37.123		
TOTAL	79	46.376		

*** p < .001
* p < .05

TABLE 17

Mean and Standard Deviation of Number of Times Prescribed Movement Function Was Used During the Second Trial

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	8.350 a (7.147)	10.600 a (6.443)	9.475 (6.813)
NON-EXPERT	11.600 a (5.195)	3.200 b (5.376)	7.400 (6.732)
TOTAL	9.975 (6.383)	6.900 (6.953)	8.438 (6.810)

Scale Values: Low values indicate less use of new movement function.
a,b: Cell means with common indices are not significantly different by the Student-Newman-Keuls procedure, p < .05.

TABLE 18

Analysis of Variance of Order of Corrections Made During the Second Trial

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	5.000	0.221	--
LEVEL OF OBSERVATION	1	26.450	1.168	< .01
EXPERTISE X OBSERVATION	1	5.000	0.221	--
RESIDUAL	76	22.651		
TOTAL	79	22.252		

TABLE 19

Mean and Standard Deviation of Number of Correction Made in the Prescribed Order by Subjects During the Second Trial

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	8.450 (5.346)	10.100 (4.564)	9.275 (4.977)
NON-EXPERT	8.450 (4.752)	9.100 (4.315)	8.775 (4.492)
TOTAL	8.450 (4.992)	9.600 (4.413)	9.025 (4.717)

Scale Values: Low values indicate less compliance with order.

TABLE 20

Analysis of Variance of Subjects' Use of New Functions for Corrections During the Second Trial

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	101.250	7.347**	0.07
LEVEL OF OBSERVATION	1	84.050	6.099*	0.06
EXPERTISE X OBSERVATION	1	4.050	0.294	--
RESIDUAL	76	13.781		
TOTAL	79	15.655		

** p < .01
* p < .05

TABLE 21

Mean and Standard Deviation of Number of Corrections Made with New Functions During the Second Trial

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	8.300 (3.404)	6.700 (4.181)	7.500 (3.850)
NON-EXPERT	6.500 (3.547)	4.000 (3.671)	5.250 (3.781)
TOTAL	7.400 (3.550)	5.350 (4.117)	6.375 (3.957)

Scale Values: Low values indicate less use of new functions.

TABLE 22

Analysis of Variance of Subjects' Use of Check Marks During the Second Trial

SOURCE	df	MS	F	w2
SUPERVISOR'S EXPERTISE	1	1.800	2.073	0.01
LEVEL OF OBSERVATION	1	3.200	3.685	0.03
EXPERTISE X OBSERVATION	1	0.800	0.921	--
RESIDUAL	76	0.868		
TOTAL	79	0.909		

TABLE 23

Mean and Standard Deviation of Values of Level of Compliance with Suggested Use of Check Marks

	LEVEL OF OBSERVATION		
SUPERVISOR'S EXPERTISE	OBSERVED	NON-OBSERVED	TOTAL
EXPERT	1.800 (1.005)	2.000 (.918)	1.900 (.913)
NON-EXPERT	1.900 (.968)	2.500 (.827)	2.200 (.939)
TOTAL	1.850 (.975)	2.250 (.899)	2.050 (.953)

Scale Values: Low values indicate greater compliance with directions.

non-expert/non-observe condition was significantly lower than the means from the expert/observed, non-expert/observed, and expert/non-observed conditions.

Note is made of the main effect for level of expertise ($F(1,76) = 7.347$, $p < .01$) for the use of the new movement function (see Table 20). Once again, this effect should be viewed critically as it is not consistent with the MANOVA effects for compliance measures.

DISCUSSION

One goal of the present study was to define more clearly the effects of perceived supervisor expertise on a subordinate's reaction to an appraisal interview with the supervisor. The manipulation check of the level of expertise variable demonstrated the success of that treatment. Subjects clearly perceived the supervisor in the "expert" condition to be more competent in skills pertaining to the task than they did the supervisor in the "non-expert" condition.

Manipulating the subjects' perception of the expertise of the supervisor clearly affected their attitudes concerning the effectiveness, accuracy, and fairness of the appraisal. On survey question numbers 5 and 10, subjects indicated that they felt the appraisal was more effective in improving their performance when they perceived the supervisor to be an expert. It seems logical to expect that an "expert" would be more likely to offer fruitful advise. Subsequently, these results are consistent with what might be considered rational expectations on the part of the subjects. Even though suggestions were identical across conditions and performances were not measurably different,

subjects demonstrated a greater confidence in the supervisor's ability to help them when he was perceived to be an "expert". The presence or absence of the supervisor during the first trial, however, made no difference in the subjects' perception of the effectiveness of the suggestions made.

Question number 8 asked subjects to evaluate the fairness and accuracy of the evaluation within a job context. Level of expertise also made a difference in the responses to this item. The expert supervisor was generally seen as providing a fairer, more accurate appraisal. Though not consistent with MANOVA effects for compliance measures in general, the interactive effect for item number 8 (see Table 13) may indicate that this level of expertise effect was more salient in the non-observed condition. Observation may have provided subjects with some reason to believe the appraisal had been fair and accurate, consequently reducing the importance of the expert/non-expert manipulation.

The present research was also intended to demonstrate the effects of differential credibility of the supervisor as determined by his presence or absence during the first trial. The question used to check this manipulation was designed to measure the subjects' perceptions of the supervisor's awareness of their performance during the first trial. The check of this treatment, however, did not result in a clear demonstration of what had been manipulated.

In fact, different treatments did occur (the supervisor was either physically present or absent during the first trial), and

the main effects resulting due to observation/non-observation conditions suggest that the two treatment conditions were perceptively different to the subjects. Unfortunately, the subjects' perception of the supervisor's knowledge of their previous performance was not, by itself, especially salient. It is possible that the presence or absence of the supervisor during the first trial may have resulted in differences in such variables as perceived pressure to perform, differential feelings of accountability for performance, perceived involvement or caring on the part of the supervisor, or other "treatments" that were, unfortunately, not measured. Consequently, interpretation of the results due to the observation manipulation are limited in that it is not known exactly what difference the observation/non-observation manipulation was making to the subjects.

Also, post-hoc analysis of the manipulation check for level of observation demonstrated how the expertise manipulation interacted with the observation manipulation. The manipulation as measured produced a homogeneous subset of cell means for the expert/observed, non-expert/observed, and expert/non-observed conditions that were all significantly higher than the non-expert/non-observed condition. Being perceived as an expert apparently enabled the supervisor to transcend the transgression of not having attended the first trial. He was given credit for somehow knowing what had occurred--even though he was not present to witness it.

In retrospect, perceiving credibility in the observations made by the expert in the absence of previous observation may not be so illogical. It may, in fact, be reasonable to assume that

an expert would be aware of the types of errors most frequently made on the task that he/she is supervising. If so, it would be logical for subjects to assume that the comments made during the appraisal session were credible--even though the supervisor had not witnessed their specific performance.

Despite the failure to document the effects of the observation treatment, the results indicated that observation prior to the appraisal affected subordinates' willingness to comply with directives given during the appraisal. For the five measures of subject compliance to the supervisor's suggestions, the most consistent effect measured was a main effect for level of observation. Three of the five (use of new functions during the practice session, use of the new movement function during the second trial, and use of the new correction functions during the second trial) were significantly higher when the subject was observed during the first trial. This was the only significant effect related to the number of new functions used during the unobserved practice session.

The measure of subjects' use of the prescribed movement function also resulted in a powerful interaction effect. Post-hoc analysis for the movement function measure revealed that the cell means for the expert/observed, non-expert/observed, and expert/non-observed conditions were not significantly different from each other. This appears to suggest that given either or both previous observation or an expert supervisor, compliance to suggestions given during the appraisal were not measurably different. Only when neither an expert supervisor nor previous

observation was given did the compliance of the subjects fall appreciatively lower than any of the other cells.

The effect mentioned above is consistent with the apparent overpowering effect of perceived expertise that was revealed by the check of the observation manipulation. The "expert" was assumed to be aware of previous performance whether present during the first trial or not. Consequently, if compliance is contingent upon perceived awareness of previous performance, compliance to directives in the expert/non-observed condition should be expected.

It was hoped that the suggestions made during the appraisal would enable those who followed them to improve their performance for the second trial. The failure of either of the dependent performance measures (time and accuracy) to result in any significant results in reference to expertise and observation--even though subjects were demonstrating different degrees of compliance to the directives--indicated that the information given was apparently not that helpful to the subjects. Observations of the subjects' efforts to implement the new functions made it clear that the added directions may have created an overload and resulted in negative transfer rather than quicker, more accurate wordprocessing. Subjects appeared to spend more time rereading the instructions for the new functions and committed many errors due to misuse of the newly introduced functions. Perhaps given a longer time frame than this experiment allowed for, compliance to the supervisor's instructions might have produced measurable results.

In any case, ultimate outcomes to any appraisal interview

will always be dependent upon the quality of the information provided to the subordinate. In this experiment, the demonstration of different levels of behavioral compliance to the supervisor's directives is the most crucial result. Given that there is any meaningful information to relay to the employee that might help him/her to improve his/her performance, compliance to the suggestions made would logically lead to differences in bottom-line performance criterion.

LIMITATIONS AND FUTURE RESEARCH

A proposed advantage of the present research over previous research in the area was the introduction of the appraisal interview question into a laboratory setting. This enabled greater control of the treatments and stronger conclusions concerning cause/effect relationships. It also, however, removed the investigation from the reality of a genuine job situation and interactions between those actually functioning in supervisory and subordinate roles. It is also true that the student population utilized as subjects did not perfectly represent the work force population for whom the results would be important. In response to the probable contention that results produced in this artificial environment may not be generalizeable to the real world, it must be conceded that what occurs in the lab is never the final word on what is occurring in the real world. The following points may, however, be offered in defense of the validity of these findings:

1. The task implemented (wordprocessing) was realistic in reference to learning and behavior that might be expected in a

real job.

2. Observations of subjects performing the task strongly indicated a generally intense desire to succeed at the task (as might be expected in a real job).

3. The superior/inferior relationship that is known to exist between an experimenter/subject in lab situations is a fair approximation of the supervisor/subordinate relationship occurring in job contexts.

4. It is known that most students have, at one time, been employed. Most have experienced some form of work evaluation (in fact, all of them if schoolwork is included). It therefore seems unreasonable to expect that their responses to the treatments given would differ markedly from any other work population.

It is also clear that the time frame of the experiment and the depth of the relationship existing between the experimenter/subject do not accurately reflect those variables as they exist in a real world setting. In response, it might be contested that if the effects measured here can be produced in such a short time span with such superficial interaction, it might be logical to expect that a real job situation would result in even stronger effects of the same nature. Since greater amounts of information can be collected over time through many interpersonal interactions of more meaning and consequence, subordinates might be expected to be more confident in their perceptions of the supervisor's expertise and credibility, and consequently more likely to act upon them.

For future research, there is a need to define more precisely the effects that observation/non-observation treatments

have on subjects. A more comprehensive set of dependent measures would have strengthened this study's conclusions concerning this question. The interaction of the a supervisor's expertise with credibility as determined by previous observation of performance also has potential for further investigation. Interviews with subordinates might help determine whether compliance to the expert's directives occur due to an assumption that he/she is aware of previous performance (perhaps because of previous experience with other performers), or whether the compliance is simply out of respect for the supervisor's "expert" status.

There may also be value in investigating the effects of a supervisor's power in terms of his/her ability to reward and punish subordinates. In reference to the dependent measures investigated in this study, in might be logical to hypothesize that such power would have a greater effect upon subordinates' compliance with directives than it would upon their perception of the value of the observations and suggestion made during the appraisal.

The ultimate environment for future research would be an actual job setting in which the researcher had the power to manipulate the type of interactions that occurred between supervisors and their subordinates. Since such intrusive arrangements are not likely to be accepted by any organization (and would probably be impossible to control if they were), a lab setting may continue to be the most practical. However, a resourceful organizational researcher might be able to take advantage of naturally occurring changes within organizations.

Time series designs over adjustments in observational strategies or changes in supervisory expertise would provide some evidence of the effects of these variables in subordinate reactions to appraisal interviews. In a work setting where supervisors are rotated, it might also be possible to identify the supervisors' observational practices and his/her level of expertise. The researcher could then measure potential changes in subordinate responses to the different types of supervisors.

CONCLUSIONS

True to the original hypotheses, these results support the contention that both supervisor expertise and observation make a difference in the way subordinates perceive and act upon information given in appraisal interviews. The following patterns are indicated by this study:

1. A high level of supervisor expertise positively affects the attitudes of subjects concerning the value of the appraisal suggestions, the appraisal's effect upon their subsequent performance, and the accuracy and fairness of the appraisal.

2. Observation prior to appraisal increases compliance with appraisal directives.

3. A high level of supervisory expertise appears to make the variable of previous observation less important. In reference to getting subordinates to comply with evaluative directives, observation prior to appraisal may only be important to those supervisors whose expertise is somewhat questionable in the eyes of their subordinates.

REFERENCES

Baird, L. S. (1977). Self and superior ratings of performance: as related to self-esteem and satisfaction with supervision. Academy of Management Journal, 20, 291-300.

Burke, R. J., Weitzel, W., & Deszca, A. G. (1982). Supervisors' and subordinates' communication experiences in managing job performance: two blind wisemen describing an elephant?. Psychological Reports, 50, 527-534.

Burke, R. J., Weitzel, W., & Weir, T. (1978). Characteristics of effective employee performance review and development interviews: replication and extension. Personnel Psychology, 31, 903-919.

Burke, R. J., Weitzel, W., & Weir, T. (1980). Characteristics of effective interviews reviewing performance: do supervisors and subordinates agree? Psychological Reports, 47, 643-654.

Greller, M. M. (1975). Subordinate participation and reactions to the appraisal interview. Journal of Applied Psychology, 60, 544-549.

Henneman, R. L. & Wexley, K. N. (1983). The effects of time delay in rating and amount of information observed on performance rating accuracy. Academy of Management Journal, 26, 677-686.

Ilgen, D. R., Peterson, R. B., Martin, B. A., & Boeschen, D. A. (1981). Supervisor and subordinate reactions to performance appraisal sessions. Organizational Behavior and Human Performance, 28, 311-330.

Knight, P. A. & Saal, F. E. (1984). Effects of gender differences and selection agent expertise on leader influence and performance evaluations. Organizational Behavior and Human Performance, 34, 225-243.

Knight, P. A. & Weiss, H. M. (1980). Effects of selection agent and leader origin on leader influence and group member perceptions. Organizational Behavior and Human Performance, 26, 7-21.

Landy, F. J., Barnes, J. L., & Murphy, K. R. (1978). Correlates of perceived fairness and accuracy of performance evaluation. Journal of Applied Psychology, 63, 751-754.

Lawler, E. E. (1967). The multi-trait multi-rater approach to measuring managerial job performance. Journal of Applied Psychology, 51, 369-381.

Shrauger, J. S. (1975). Responses to evaluation as a function of initial self-perceptions. Psychological Bulletin, 82, 581-596.

Shrauger, J. S. & Lund, A. K. (1975). Self evaluations and reactions to evaluations from others. Journal of Personality, 43, 94-108.

Smircich, L. & Chesser, R. J. (1981). Superiors' and subordinates' perceptions of performance: beyond disagreement. Academy of Management Journal, 24, 198-205.

APPENDIX A

TEXT A: Eight years ago it was the frightening, unknown killer in a mystery. today, it si considered jost one of many forms of penumonia.

But researchers are stilll probing basic questions about legionnaires' disease and the bacteria that cause it.

The disease and the bacteria, "Legionella," were named for na outbreak of pneumonia in which one hundred eighty-two people in Philadelphia, most of whom had attended a state state American Legino convention there in July 1976. Twenty-nine people deid before doctors and federal health officials traced the disease to Legionella that had bene blown into the air from the air conditioning system in a convention hotel. The convention was held in Philadelphia.

TEXT B: David, the 12-year-old "Bubble Boy" who lacks immunity to disease, was removed for the first time from a germ-free unit so doctors could investigate why he has recurring fever, texas Children's Hospital authorities said ni Houston Wednesday.

Dr. William Shearer, chief of allargy and immunology srevices at Texas Children's Hospitall, said david is not in danger, but had to be removed from his normal unit os tests could be performed.

"We are conducting one hundred twenty-three tests to determine the cause of these symptoms," Shearer said said. "Because these symptoms requier close medical supervision, we cuold not effectively treat thme while he was in his isolation unit." Dr. Shearer is an immunology specialist.

APPENDIX B

CORRECTIONS FOR TEXT A

Corrections:

1. Capitalize the letter "t" in the word "today".
2. Correct the spelling of "jost" to "just".
3. Capitolize the letter "l" in "legionnaires".
4. Change "one hundred eighty-two" to the numeric form "182".

Reversals:

5. Change the incorrect word "si" to the correct word "ia".
6. Change "penumonia" to the correct spelling "pneumonia".
7. Change the incorrect word "na" to the correct word "an".
8. Change the spelling of "Legino" to the correct version "Legion".
9. Correct the spelling of "deid" to "died".
10. Correct the spelling of "bene" to "been".

Deletions:

11. Delete the third "l" in the word "stilll".
12. Delete the incorrect duplication of the word "state".
13. Delete the entire sentence "The convention was held in Philadelphia".

Insertions:

14. Insert the word "medical" between the words "a" and "mystery".
15. Insert the word "deadly" between the words "the" and "bacteria"

CORRECTIONS FOR TEXT B

Corrections:

1. Capitalize the letter "t" in the word "texas"
2. Correct the spelling of "allargy" to "allergy".
3. Capitalize the letter "d" in "david".
4. Change "one hundred twenty-three" to the numeric form "123".

Reversals:

5. Change the incorrect word "ni" to the correct word "in".

6. Change "srevices" to the correct spelling "services".

7. Change the incorrect word "os" to the correct word "so".

8. Change the spelling of "requier" to the correct version "require".

9. Correct the spelling of "cuold" to "could".

10. Correct the spelling of "thme" to "them".

Deletions:

11. Delete the second "l" in the word "Hospitall".

12. Delete the incorrect duplication of the word "said".

13. Delete the entire sentence "Dr. Shearer is an immunology specialist."

Insertions:

14. Insert the work "isolation" between the words "germ-free" and "unit".

15. Insert the word "imminent" between the words "in" and "danger".

APPENDIX C

Please respond to the following statements with a VT if the statement is very true concerning your perception of the experiment, T if it is true, N if it is neither especially true nor false, F if it is false, and VF if it is very false.

1. The person running the experiment was extremely knowledgeable in the use of microcomputers for wordprocessing. VT T N F VF

2. These particular wordprocessing functions were very easy to learn. VT T N F VF

3. It was obvious that the experimenter knew enough about what I had done on the first trial of the task to make very good suggestions about what I should try on the second. VT T N F VF

4. The directions for this word processing task were very easy to understand VT T N F VF

5. The experimenter's suggestions were all very effective. Every one clearly helped me to increase my level of performance on the second trial. VT T N F VF

6. It was a very good idea to introduce the additional functions that the experimenter told me about after the first trial. VT T N F VF

7. The observations that the experimenter made of my performance of the first trial were very accurate. VT T N F VF

8. Had this task been my job and the experimenter my boss, I would have considered the evaluation of my work very fair and accurate. VT T N F VF

9. I payed close attention to the experimenter's suggestions during the comments after the first trial. VT T N F VF

10. The experimenter's appraisal of my first trial definitely enabled me to improve my performance on the second trial. VT T N F VF

11. I made good use of the suggestions that the experimenter made during the appraisal of my first trial. VT T N F VF

Please estimate your abilities previous to this experiment by responding to the following:

12. I would describe my typing skills as:

a. very good
b. good
c. about average
d. poor
e. very poor

13. I have used this particular word processing program before.

a. true
b. false

14. I have had the following amount of experience with word processing in general previous to this experiment:

a. extensive experience
b. a good deal of experience
c. some experience
d. little experience
e. no experience whatsoever

15. The computer experience I have had has been with:

a. a microcomputer (similar to this one)
b. a mainframe computer (accessed by a video terminal or card reader)
c. both microcomputer and mainframe computer
d. not applicable--I have had no experience

16. I own a microcomputer.

a. true
b. false

APPENDIX D

MEANS AND STANDARD DEVIATIONS

	MEAN	STD DEVIATION
EXPERTISE MANIPULATION	2.338	1.079
OBSERVATION MANIPULATION	1.725	0.842
VALUE OF SUGGESTIONS	2.088	0.917
FAIRNESS AND ACCURACY	2.163	0.974
ENABLED IMPROVED PERFORMANCE	1.950	0.953
NEW FUNCTIONS USED IN PRACTICE	2.850	4.290
USE OF MOVEMENT FUNCTION	8.438	6.810
CORRECTIONS MADE IN ORDER	9.025	4.717
USE OF CORRECTION FUNCTIONS	6.375	3.957
USE OF CHECK MARKS	2.050	0.953
ERRORS ON SECOND TRIAL	4.875	10.850
TIME IN SECONDS FOR SECOND TRIAL	785.075	403.209

PRE-EXISTING DETERMINANTS OF APPRAISAL INTERVIEW SUCCESS

by

Matt L. Riggs

B.A., Bethany College, 1979

AN ABSTRACT OF A MASTER'S THESIS

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Psychology

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1986

Abstract

Supervisor observation of work performance and supervisor expertise were expected to increase the success of a subsequent appraisal interview. Success of the interview was defined by subjective reactions of the subordinates to the value of the appraisal, their subsequent behavioral compliance with directives given during the appraisal, and their ultimate performance of the same task following the appraisal. Although no effects on the performance criteria were registered, expertise improved subordinates' perceptions of the effectiveness, accuracy and fairness of the appraisal, and observation and expertise increased compliance with recommendations given during the appraisal.